

SEN CUTLER

Data Engineer | Tacoma, WA (remote) | 1 (253) 349-5499 | sxcutler@protonmail.com | <https://sen-cutler.net/>

Available for fully remote contract engagements, US-based, Pacific time

PROFESSIONAL SUMMARY

Independent data engineer specializing in API integrations, ETL pipelines, backend automation, and platform-to-platform data migrations. Built and productionized a reusable async migration toolkit now used across nine client onboardings. Python throughout, with PostgreSQL, AWS, and Docker. AI/ML work spanning LLM training data, semantic matching and ranking, and OCR-heavy document extraction.

EXPERIENCE

Data Engineer — Self-employed, Remote — April 2024 to Present

- Built a reusable async Python migration toolkit (~8,800 SLOC) against the Lawmatics API for a legal-tech client; used it to execute nine customer onboarding migrations from legacy legal CRM software.
- Built an ETL framework for repeated ingestion from external APIs into PostgreSQL, with shared extract, transform, validate, and load stages and resumable runs.
- Matched and de-duplicated a 10.8M row dataset in PostgreSQL, tuning name, company, and location matching logic to reduce both false positives and false negatives.
- Reduced runtime of a Python/MySQL text-segmentation job across 2.8M rows from 19 hours to 2.5 hours over successive optimization passes.
- Built a 1.1M token AI training corpus from 13 book-length PDFs, handling widespread OCR errors, then used the OpenAI API to generate question-answer pairs for every chunk.
- Deployed containerized backend services to AWS App Runner and Lambda, covered by unit and end-to-end API tests.
- Built ML-driven semantic matching and ranking engines over unstructured text, blending embeddings, BM25, and latent semantic analysis with configurable weighting and negative-example exclusion.
- Extracted structured data from difficult sources at scale — 107 MB of XML, 117 PDF book scans of uneven quality, a 190 MB JSON chat-log export — into PostgreSQL and client-ready spreadsheet deliverables.

TECHNICAL SKILLS

Languages & Core	Python (asyncio, pandas, pytest), SQL, data wrangling
AI & ML	Embeddings, vector search, OpenAI API, prompt engineering, preparing training data, OCR, clustering, grid search, latent semantic analysis
Data stores	PostgreSQL, MySQL, SQLite, MongoDB, vector databases
APIs	REST integration and development, FastAPI, webhook intake, OAuth
Cloud & Infrastructure	AWS (App Runner, Lambda), Docker, Google Cloud Platform
Platforms	Google Sheets API, Gmail API, Lawmatics API

CERTIFICATIONS

- IBM AI Engineering Professional Certificate (V3) — IBM, December 2024
- IBM Machine Learning Professional Certificate — IBM, December 2024
- Docker Mastery: with Kubernetes + Swarm — Udemy, October 2023

CREDENTIALS

- Top Rated+ on Upwork with a 100% Job Success Score: <https://www.upwork.com/freelancers/kernef>
- Letter of recommendation: <https://sen-cutler.net/recommendation/>

EDUCATION

B.S. in Statistics, University of Washington, Seattle, WA